Designing Visuo-Haptic Illusions with Proxies in Virtual Reality: Exploration of Grasp, Movement Trajectory and Object Mass

Martin Feick  
DFKI, Saarland Informatics Campus  
Saarbrücken, Saarland, Germany  
martin.feick@dfki.de

Kora Persephone Regitz  
DFKI, Saarland Informatics Campus  
Saarbrücken, Saarland, Germany  
kora.regitz@dfki.de

Anthony Tang  
University of Toronto, Faculty of Information  
Toronto, Ontario, Canada  
tonytang@utoronto.ca

Antonio Krüger  
DFKI, Saarland Informatics Campus  
Saarbrücken, Saarland, Germany  
antonio.krüger@dfki.de

Figure 1: A user moves a proxy whisk inside a proxy pot (left). By offsetting the virtual hand/tool from the real hand/tool we can simulate virtual pots of different sizes providing realistic haptic sensations when stirring. Our studies investigate the extent to which we can use such illusions by exploring the potential effects of grasp, movement trajectory and object mass on the discrepancy which can be introduced while remaining unnoticed by a user.

ABSTRACT

Visuo-haptic illusions are a method to expand proxy-based interactions in VR by introducing unnoticeable discrepancies between the virtual and real world. Yet how different design variables affect the illusions with proxies is still unclear. To unpack a subset of variables, we conducted two user studies with 48 participants to explore the impact of (1) different grasping types and movement trajectories, and (2) different grasping types and object masses on the discrepancy which may be introduced. Our Bayes analysis suggests that grasping types and object masses (≤ 500 g) did not noticeably affect the discrepancy, but for movement trajectory, results were inconclusive. Further, we identified a significant difference between (un)restricted movement trajectories. Our data shows considerable differences in participants’ proprioceptive accuracy, which seem to correlate with their prior VR experience. Finally, we illustrate the impact of our key findings on the visuo-haptic illusion design process by showcasing a new design workflow.

CCS CONCEPTS
- Human-centered computing → Virtual reality.

KEYWORDS
Visuo-Haptic Illusions, Grasp, Movement Trajectory, Object Mass

ACM Reference Format:

INTRODUCTION

Virtual Reality (VR) has become a consumer technology in recent years; however, our understanding of haptic feedback remains underdeveloped. Many studies have shown that providing adequate
haptic feedback for objects and the environment significantly enhances the experience [24, 31, 39–41, 54, 79]. Yet it remains challenging to render haptic feedback to correspond exactly with the virtual experience; properties such as shape and size [24, 41, 63, 79], texture [21, 73], weight [17, 63, 76], and functional object parts [24, 81] are very challenging to replicate. One promising approach is the use of so-called proxies – physical real-world objects that resemble parts of, or even entire, virtual objects [41]. For example, a sphere can act as a “stand-in” for a virtual globe. Ideally, a single physical proxy can stand in for multiple virtual objects. To achieve this, researchers have proposed the use of visuo-haptic illusions that leverage the visual dominance phenomenon, where vision usually overrides proprioception when the two senses conflict [14, 15]. This method introduces an unnoticeable offset between the virtual and real world such as by mapping haptic features in the environment [7, 48]. Several research projects demonstrate that this technique can be successfully used to enhance proxy-based interactions [1, 6, 7, 11, 26, 38, 48, 61, 68], overcome a limited interaction space [5, 9, 16, 23, 66, 77] or improve a VR controller’s capabilities [75]. An ever-present question is the extent to which we can introduce differences between the real and virtual world without creating a semantic violation (i.e., where the discrepancy is too large and can be noticed by a user). To address this, researchers have conducted psychophysical experiments reporting detection thresholds for specific use cases and scenarios [1, 9, 11, 23, 26, 34, 66, 70, 77]. However, because these are not a systematic investigation on how illusions can be incorporated beyond specific use cases, we do not have a generalizable design approach on how to incorporate visuo-haptic illusions in VR.

Imagine the following scenario: Lisa is a VR designer and needs to design a virtual cooking class application for remote students. She wants to include haptic feedback, because she knows that the acquired skills transfer better to the real world when the experience is both visually and haptically sufficiently well rendered [10, 41, 50]. Lisa does not know which proxies the students can access, but aims to be as inclusive as possible by requiring only a minimal set of proxies in order to participate. For instance, a student might only have a single proxy pot which matches the size of one virtual pot in the application (see Figure 1 B). By incorporating illusions she allows the students to use a single proxy pot for various virtual pots of different sizes (see Figure 1 A and C) creating realistic sensations while stirring with a proxy whisk. However, when using different kitchen utensils, each tool comes with its own properties and requirements, i.e., how it can be moved, how heavy it is, how it should be handled and so on. She asks herself, “Can I use the same thresholds for all the different interactions and available proxies or do these somehow affect the discrepancy which can be introduced?” Variables such as movement trajectory, distance, grasping type, speed, time, properties of the object (shape, size, mass etc.) as well as complexity of the application (distracting factors) might play key roles. However, we do not know the extent to which these variables limit or even extend the illusion space.

In this work, we aim to untangle the contributing factors leading to a semantic violation by isolating the potential effects of: grasping types (proxy shape and size), movement trajectories and object masses (≤ 500 g). To do so, we conducted two user studies with 48 participants, determining their detection thresholds for the different conditions and comparing them. Our traditional frequentist analysis did not reveal a significant effect of the variables grasping types, object masses and movement conditions (linear and circular) on the amount of disparity which can be introduced. The computed Bayes factors suggest that for movement trajectory, this was due to insensitive data. For grasping type and object mass, there was evidence for the absence of an effect on the thresholds given the data. In contrast, we found a significant difference for movement trajectory when comparing linear restricted vs. unrestricted movement, showing that proxies that limit the degrees of freedom (e.g., 1D slider) allow for greater offsets. Further, our descriptive analyses revealed that there are substantial per-user differences in human proprioceptive acuity, which seem to correlate with their prior VR experience. Finally, we illustrate the impact on the design process of visuo-haptic illusions by applying our results to Lisa’s design scenario.

Our eventual goal is to provide a systematic and generalizable approach to include visuo-haptic illusions in the design process of new VR experiences. In this work, we make four contributions:

1) Report our estimates for the conservative detection thresholds for all study conditions.
2) Unpack the impact of grasping type, manipulation trajectory and proxy mass on visuo-proprioceptive conflicts and investigate their differences.
3) Provide initial design guidelines for incorporating visuo-haptic illusions into the design workflow demonstrated through our design example.
4) Provide evidence of the differences in humans’ sensitivity to illusions in VR.

2 RELATED WORK

In this work, we make the first attempt to untangle the proprioceptive factors contributing to the successful design of visuo-haptic illusions in VR – which remain unnoticeable by a user when manipulating proxies in VR. To do so, we outline some of the influential work in the field of haptics in VR, specifically in the context of proxies. Next, we discuss the nature of visuo-haptic illusions and finally, we look at how humans grasp and manipulate virtual objects embodied by a physical proxy.

2.1 Haptics & Physical Proxies

In the context of haptics in VR, we broadly distinguish between two types of haptic feedback, active and passive. The latter relies on physical properties such as size and shape [24, 40, 79, 81], textures [21, 73] etc., where a real-world object is mapped to a virtual counterpart [40, 41]. These physical objects are usually referred to as proxy objects which act as “stand-ins” for multiple virtual objects. One challenge is finding suitable physical proxies that match the virtual objects [55]. To address this, researchers have explored several creative approaches on how to obtain an optimal proxy object: by searching the user’s environment for the optimal proxy [39, 64], or using self-assembling robotic devices [79], reconfigurable devices [81], or modular toolkits [24]. In contrast, active haptic feedback utilizes computer-generated actuation to present a haptic stimulus.
to users. A prominent example is the Phantom haptic device allowing users to feel the virtual environment and its boundaries [52]. Many other prototypes provide force feedback when a user grasps or touches virtual objects [18, 19, 49, 65] and moreover, enable them to feel textures [73], weight [17, 37, 76] and stiffness [60, 69]. Another class of haptic devices are encounter-type systems [53]. They use actuation to position themselves in the environment, providing haptic feedback such as for touching and grasping objects [2, 4]. A common method to expand haptic devices’ capabilities and to enrich interactions is to include visuo-haptic illusions.

2.2 Visual Dominance & Visuo-haptic Illusions

When humans encounter a situation where there is a sensory conflict between visual and another sensory modality, humans rely more on visual information to resolve the conflict [15, 32]. This is of great interest to the VR community, since when there is a discrepancy between the real world and the virtual world, designers can effectively rely on synthetic visual information to override other sensorial input [14, 35]. For instance, researchers have leveraged this effect to redirect walking by subtle warping of the virtual space [66]. Similarly, one can redirect a user’s hand by offsetting the virtual hand from the real hand [77]. Users unconsciously compensate for this, resulting in them visually touching different objects while in fact, they have been redirected to the same physical proxy [5, 16]. This method is called haptic retargeting, and has also been applied to two-handed (bimanual) interactions [34].

Visuo-haptic illusions exploit the visual-dominance effect, allowing designers to use physical proxies that only share some attributes with the virtual objects [6]. For example, Kohli [48] presents a redirected touch technique, warping the virtual space, allowing a single physical proxy to act as a proxy object for multiple virtual objects with different geometries. Ban et al. [7] develop a perception-based shape display using a simple cylinder primitive with a haptic bump allowing them to display various shapes. Bergström et al. [11] change the perceived object size by morphing a human’s virtual hand, while Samad et al. [61] change the perceived weight by applying Control-Display ratio (C/D) manipulations. Changing the C/D ratio leads to larger (C/D > 1.0) or smaller (C/D < 1.0) virtual movements than physically performed. Tinguy et al. [70] investigate how different a physical proxy can be with respect to its width, local orientation, and curvature, reporting estimates for detection thresholds. Feick et al. [26] look at how much discrepancy between proxies’ manipulable parts can be introduced and found that movable object parts can have quite substantial differences.

Illusions can also be used to overcome limitations of haptic devices or to expand their interaction space. For instance, VR Grabbers is a chopstick-like passive VR controller enabling users to precisely select and manipulate virtual objects, and by introducing a positional offset even allows them to grab objects beyond its hardware capabilities [75]. Abtahi and Follmer [1] use visuo-haptic illusions to overcome existing limitations of shape displays. Gonzalez et al. [33] propose dynamic visuo-haptic redirection to compensate for the workspace limits and device latency issues of encounter-type haptic devices. PseudoBend [38] is a proof-of-concept prototype creating the sensation of twisting, stretching, and bending a stiff bar by combining visual feedback with vibrotactile stimuli. GamesBond [59] is a bimanual controller and utilizes haptic illusions to simulate physically connected objects (e.g., a rope). Zenner et al. [78] use a weight-shifting proxy and haptic retargeting to simulate weight shifts beyond the proxy’s capabilities. Strandholt et al. [68] apply positional offsets between virtual and proxy tools, e.g., a hammer, to provide the sensation of manipulating a second proxy (i.e., a nail). An unexplored but potentially crucial aspect is how humans grasp or hold the object. Requirements might change if users hold a hammer with a wrapped whole hand grip, versus a small nail with just two fingers. Understanding how grasping types might affect visuo-haptic illusions is the focus of our work.

2.3 Grasping in VR

Human grasping is a complex and necessary interaction to leverage the full potential of objects in our environment. For instance, there are various ways to grasp a simple mug [28], and grasping becomes even more complex when looking at interactions such as opening a bottle, where humans seamlessly transition between several grasping types [20] – demonstrating the dynamics and unpredictability of these interactions. Grasping objects in a virtual environment is one of the core interactions with virtual objects. Therefore, a large body of work has looked at ways we can haptically support this complex interaction in VR. Solutions range from dedicated controller which support grabbing objects of different geometries [18, 75] using pinching-type gestures [17, 49, 65], to matching proxies to provide compelling grasp sensations [39], to high-fidelity haptic rendering using exoskeletons [36]. Yet it is unknown how differences in grasping poses may affect the effectiveness of visuo-haptic illusions. Thus far, researchers have mostly limited the types of interaction that can be performed, such as a simple touch [7]. Prior work shows that we can use visuo-haptic illusions in specific, bespoke situations and implementations. In this work, we take a first step towards developing a generalizable understanding of the extent to which different proprioceptive factors lead to semantic violations. Our eventual goal is to provide a set of design guidelines that can be applied to existing as well as new VR experiences, allowing developers to easily incorporate visuo-haptic illusions into their workflow. Here, we assume that the touch points between virtual and physical object are correctly rendered.

3 IMPACT OF GRASP, OBJECT MASS AND MOVEMENT TRAJECTORY

There are several variables which may contribute to a semantic violation. In this work, we explore the effect of three such variables: grasping type, object mass, and movement trajectory. In this section, we outline our selection process and discuss why it is important to understand the impact of these variables to develop a generalizable design approach for visuo-haptic illusions.

3.1 Does a user hold an object affect how much discrepancy may be introduced?

Humans choose the correct grasping type based on the underlying task requirements [13, 20, 27] and objects’ characteristics [28] (particularly, the shape of the object [20]). These variables remain entangled and can therefore only be considered holistically. Cutkosky et al.’s [20] grasping taxonomy broadly distinguishes
between power, (intermediate) and precision grasps. As the name already suggests, power grasps are used to manipulate heavier/larger objects or when dexterity is secondary. On the other hand, precision grasps are primarily for fine grained manipulations. Hence, different muscle groups are involved when changing or adjusting the grasping type [67] which motivates the question of whether the grasping pose itself affects how much discrepancy between the real and virtual world may be introduced. Below, we outline the selection process of the four grasping types we chose for our studies.

We analyzed several grasping taxonomies describing between 14 and 33 grasping types by comparing their similarities and differences [20, 29, 44, 67]. Our goal was to include one representative grasping type per established category across the different taxonomies, maximizing the likelihood to identify potential differences. To do so, we prioritized the grasping types in each category based on their usage frequency in four different application areas: housekeeping, machinery, food preparation, and laundry [28, 80]. We selected the four grasping types for our study according to their usage frequency, kinematic differences [67, 74], and distinct object characteristics (i.e., size and mass) to obtain a diverse set of grasping types and to increase variability. This also aligns with the proposed optimal grasp set by Feix et al. [28]. We designed the corresponding proxy grasps (see Figure 2 and Table 1) based on Feix’s grasp size analysis with real-world use cases in mind.

### 3.3 Does how the proxy is moved affect how much discrepancy may be introduced?

Findings in the hand redirection domain show that the movement direction and distance with respect to the user’s body significantly affect the detection thresholds [9, 26]. For instance, distancing one’s hand from one’s body allows for much greater discrepancy than vice versa (i.e., bringing one’s hand closer to oneself) [9]. We included two distinct manipulation trajectories to explore if possible differences between grasping types occur for different movement directions. Rather than opting for the three main axes and limiting the degrees of freedom to only one, we used Lissajous-figures, which are used in the motor learning field [8], to systematically include more complex and rich interactions. We chose a 1:1 frequency ratio and 0° phase offset, and a 1:1 frequency ratio and 90° phase offset, resulting in a linear and circular movement (Figure 3). The furthest waypoint was set to 30 cm [77] to ensure that participants could physically reach it without fully extending their arm, which would provide a strong proprioceptive cue. To be able to compare the two movement trajectories, we used the distance point D in relation to a user’s torso and mapped the C/D-ratio intervals for physical movements and their corresponding virtual representation to one another. As a result, the only difference between the two trajectories is the total movement distance covered (i.e., circle perimeter $u = \pi \times D > \text{linear distance D}$).

### 3.4 C/D-ratio Manipulations

A common approach to create illusions in VR is to manipulate the C/D-ratio, exploiting the visual dominance phenomenon [26, 61].

Figure 2: The four grasping types—Lateral (A), Medium Wrap (B), Tripod (C) and Writing Tripod (D)—that we chose following the selection process described above. The objects were designed and 3D printed according to the grasp requirements.
Table 1: Grasp classification, correlated object masses and dimensions, and examples for our four grasping types.

<table>
<thead>
<tr>
<th>Grasping Type</th>
<th>Lateral</th>
<th>Medium Wrap</th>
<th>Tripod</th>
<th>Writing Tripod</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grasp class [20, 67]</td>
<td>intermediate and flat</td>
<td>power and cylindrical</td>
<td>precision and spherical</td>
<td>precision and distal</td>
</tr>
<tr>
<td>Mass (avg.) [27, 28]</td>
<td>150 g</td>
<td>400 g</td>
<td>150 g</td>
<td>20 g</td>
</tr>
<tr>
<td>Dimensions [27, 28]</td>
<td>up to 2 cm thick</td>
<td>4.5 cm in diameter</td>
<td>3 cm in diameter</td>
<td>1 cm in diameter, tilt angle of 62.4° [74]</td>
</tr>
<tr>
<td>Examples [13, 27, 28, 80]</td>
<td>towels, keys, paper, mug handle, cards</td>
<td>bottles, cans, vacuum, mop, handles</td>
<td>doorknobs, salt/pepper shaker, chess pieces</td>
<td>drawing and writing tasks, kitchen and workshop tools</td>
</tr>
</tbody>
</table>

We utilize this method to scale up a user’s real-world movement, resulting in a larger virtual movement than physically performed, which can be done by introducing a gain factor. Figure 6 illustrates the effect on both movement trajectories, linear and circular. In this work, we are primarily interested in expanding the interaction space and thus, we only consider C/D gains ≥ 1.0 resulting in larger virtual movements than physically performed.

4 EXPERIMENTS

The objective of this work is to understand how variables contribute to a semantic violation, which provides a better understanding of how visuo-haptic illusions may be used to enhance VR experiences. We designed two studies investigating: (1) effects of grasping type and the manipulation trajectory, and (2) the effect of different grasping types and object mass (see Figure 4). To do so, we varied the discrepancy between the physical proxy and virtual object position by applying different C/D-ratios for simple manipulation tasks. Our main research objectives were:

R1: Does a user hold the object affect how much discrepancy may be introduced?

R2: Does the movement trajectory affect how much discrepancy may be introduced?

R3: Does how heavy an object is affect how much discrepancy may be introduced?

R4: Does performing restricted vs. unrestricted movements affect proprioceptive accuracy?

R5: Do participants differ in their proprioceptive acuity?

We conducted two psychophysical threshold experiments to investigate the effects on the Conservative Detection Thresholds (CDT) [26, 78] for different independent variables outlined in studies 1 and 2 below. Both studies were executed at the same lab facility and used the same simple virtual environment consisting of two tables, the experimental setup, and an instruction screen. Participants remained seated on a chair throughout the experiment and were carefully positioned in front of the physical setup. Participants wore an HMD with their dominant hand being tracked. They were told to manipulate the proxy until it matched a target position displayed in the virtual word. After they successfully established the position, an alternative forced-choice (‘yes’ or ‘no’) question appeared, and they were asked whether they noticed a manipulation or not [66]. In the linear movement condition (study 1 and 2), they responded to: “My virtual hand moved in a wider circle than my own” [26]. In contrast, in the circular condition (study 1) they responded to: “My virtual hand moved faster than my own”. In both studies, participants were informed about the procedure, and we explicitly showed them the effect of C/D-ratio manipulations multiple times during the warm-up phase. They were told to report a manipulation as soon as they noticed it, thus targeting the most conservative case.

4.1 Study 1: Effects of Grasping Type and Movement Trajectory

In study 1, we compared the four grasping types (medium wrap, lateral, tripod and writing tripod) across two restricted movement trajectories (linear and circular manipulation). We used two different physical setups enabling us to restrict a user’s movement, preventing involuntary path deviations and neglecting an object’s weight. This allowed us to isolate the effects that different grasping types and movement trajectories may have on the perception. Through this study, we wanted to understand whether we could use the same thresholds when using different grasping postures and manipulating proxies along different trajectories.

4.1.1 Design.

We utilized an adaptive psychophysical 1-up-1-down interleaved staircase procedure with a 4x2 within-subjects design. We had two independent variables: GRASPING TYPE (lateral vs. medium wrap vs. tripod vs. writing tripod) x MOVEMENT TRAJECTORY (linear vs. circular). In total we investigated 8 conditions which were counterbalanced using a Latin square (n=8).

We used a 1-up-1-down interleaved staircase procedure exposing participants to different stimuli (C/D-ratios) repeatedly. Using a fixed step size, we target the Conservative Detection Threshold (CDT) or point of subjective equality [43, 47]. The interleaved staircase uses a descending and an ascending sequence, and randomly assigns the next trial to one of the sequences. The procedure increases the next stimulus if a participant fails to detect the current stimulus and decreases the next stimulus if the user detects the manipulation. A directional change within a sequence is noted as a reversal point. We used the number of reversal points (r=5) as a convergence criterion. Based on previous studies in this field, we chose 1.0 and 2.0 for our range of manipulation factors with a 0.1 fixed step size [9, 26, 77]. Following our pilot tests, we selected 1.0 († asc.) and 1.8 (↓ desc.) as the starting values for the procedure to allow for quicker convergence.
4.1.2 Participants.
We recruited 24 right-handed participants (eleven females, thirteen males), aged 20–36 (mean: 26.42; SD: 3.65) from the general public and the local university. Participants had a range of different educational and professional backgrounds including media informatics, computer science, education, pharmacy, angiology, neuroengineering, embedded systems, data science and artificial intelligence. All participants reported normal or corrected-to-normal vision and did not report any known health issues which might impair their perception or proprioception. Eight participants had never used VR before, ten had used it a few times (one to five times a year), no one reported using it often (6–10 times a year), and six others used it on a regular basis (more than 10 times a year). Ten participants reported that they had not played VR games before, nine people responded sometimes or infrequently (1–5 times a year), one often (6–10 times a year), and four people on a regular basis (more than 10 times a year). Participants not associated with our institution received €10 as remuneration for taking part in the experiment. The study was approved by the University’s Ethics and Hygiene Board.

4.1.3 Apparatus.
In study 1, we used the apparatus shown in Figure 5, consisting of an HTC VIVE\(^1\) Pro Eye tracking system and an Optitrack\(^2\) system with five Flex13 cameras. On the software side, we used SteamVR\(^3\) (v 1.1.7), OpenVR SDK\(^4\) (v. 1.1.4) and Motive\(^5\) (v. 2.3.0) for motion capturing and running a simple virtual scene, which was developed in Unity3D\(^6\) (v. 2020.2.1f1) and was executed on an Acer Predator Orion 5000 POS-615s offering an Intel® Core i9 10900k CPU, 32 GB RAM and an Nvidia® GeForce RTX 3080. To support the initial grasping phase, we included hand tracking through a Leap Motion\(^7\) controller (core v. 4.5.0) using an androgynous hand representation without noticeable characteristics as suggested by Schwind et al. [62] to prevent unwanted effects [56]. We built two different physical setups allowing us to restrict users’ movement. For the linear setup, an 80 cm camera slider was used, forcing participants to translate the proxy alongside its path, thus, not allowing any path deviations [26]. Additionally, this mechanism enables us to ignore object mass. A custom mount was 3D printed allowing us to quickly swap out the objects for the different study conditions. The circular setup makes use of a lazy Susan turntable (metal bearing) with a laser-cut wooden plate and a custom mount which: (1) could rotate around its center using a second bearing, and (2) hosted the magnetic mount for attaching the different study objects. The two setups were fixed on tables and therefore could not be accidentally moved by our participants.

The four objects were 3D printed using PLA, and included 3D printed conductive parts (composite PLA – Electrically Conductive Graphite) to enable touch sensing. Following Tinguy et al. [71], we used a combination of optical tracking and capacitive sensors to improve the visuo-haptic synchronization and immersion in VR. In addition, by snapping the virtual hand to the virtual object when physically touching the proxy, we could avoid hand tracking issues [26]. For touch sensing we used an Arduino Uno running a capacitive touch sensing sketch — transmitting (no-)touch events to the Unity3D program through serial port communication. The experimental logic was implemented using the Unity Experiment Framework (UXF v.2.1.1) [12] and the Unity Staircase Procedure Toolkit\(^8\).

4.1.4 Procedure.
After a general introduction to the study, informed consent and explaining the hygiene measures in place, participants filled in the demographics questionnaire. Following this, they were introduced to VR, the system, and the task. Participants were guided through an open-ended practice round to familiarize themselves with the task and the system. In the second step, we exposed them to trials with and without manipulation factors to illustrate the effect, and only proceeded once they felt confident in detecting a manipulation.

Participants were instructed to grasp the proxy object as indicated

---

\(^1\)https://www.vive.com/
\(^2\)https://optitrack.com/
\(^3\)https://www.steamvr.com/en/
\(^4\)https://github.com/ValveSoftware/openvr
\(^5\)https://www.optitrack.com/software/motive/
\(^6\)https://unity.com/
\(^7\)https://developer.leapmotion.com/
\(^8\)https://github.com/AndreZennert/staircase-procedure/
and to maintain the pose through each round of the experiment. The experimenter ensured that participants did not change their grasping pose unintentionally. They were told to move the object to the target position with a consistent and comfortable speed. The system monitored that they stayed within a reasonable time limit. Once they reached the goal position, the forced-choice question appeared, and the object needed to stay within a 5mm distance for the question to remain visible. Participants were instructed to respond to the question as quickly as possible by pointing to either ‘yes’ (there was a manipulation) or ‘no’ (there was no manipulation) using the VIVE controller in their non-dominant hand. In our pilot experiments, we observed that participants carried a bias from the previous staircase round to the next. To address this and cope with proprioceptive fatigue [58], participants took a longer break after each staircase round (by removing their headset). Before starting a new round, participants were given five calibration trials with no manipulation factor, helping them to “re-calibrate” themselves. After completing the eight conditions, participants filled in a Simulator Sickness Questionnaire (SSQ) [45]. The total experiment took about 60–70 min.

4.1.5 Data Collection.
We collected data from five sources: a pre-study questionnaire for demographic information; the subjective responses to the forced-choice staircase question; system logs (including trial times, object position and orientation, and velocity using UXF [12]); field notes and observations; and a poststudy SSQ in VR using the VRQuestionnaireToolkit [25].

4.1.6 Results.
We report our estimates for the conservative detection thresholds using different grasping types (lateral, writing tripod, medium wrap, and tripod) along two restricted movement trajectories (linear and circular). Then, we analyze the results with respect to our research objectives.

Detection Thresholds for Grasping Types and Movement Trajectories. We collected 4346 responses through the interleaved-staircase procedure. On average, it took participants 22.6 (SD: 3.6) trials to reach convergence. For each participant, we obtained eight thresholds (i.e., one per condition) by averaging the last four reversal points within the ascending and descending staircase sequence. The overall thresholds for the eight study conditions were determined by computing the mean across all 24 individual threshold values [26, 78]. The results can be found in Figure 7. All 192 staircase plots from study 1 are available in the appendix. Our analysis from the SSQ responses shows an increased Total Severity (TS) score, mean = 21.04, SD = 12.02 (P10 and P21 SSQ data lost). We hypothesize that this was a result of participants wearing a medical mask under the
headset (as a COVID-19 hygiene measure) which increased sweating and discomfort according to participants’ post-study comments.

**Analysis.** We statistically analyzed our data using a Two-Way Repeated Measures ANOVA on the two independent variables, movement trajectory with two levels and grasping type with four levels. First, we identified two significant outliers using the box plot method, which we removed from the dataset for the analysis step. The dataset met the normality assumptions at α = .05, verified through a Shapiro-Wilk test. We checked the assumption of sphericity using Mauchly’s test and applied Greenhouse-Geisser corrections to the within-subject factor grasping type, because sphericity was violated. To further investigate our data, we conducted a Bayesian ANOVA using the BayesFactors R package with default priors (v. 0.9.12-4.3). Effects are reported as the Bayes factor for the exclusion of a particular effect ($BF_{excl}$), calculated as the ratio between the likelihood of the data given the model with the effect vs. the next simpler model without that effect [46].

**Research Objectives R1 & R2.** The Two-Way Repeated Measures ANOVA did not reveal a main effect on both variables, manipulation trajectory ($F_{1,21} = 2.292, p = .145$) and grasping type ($F_{3,63} = .928, p = .427$), within our collected data. There was also no interaction effect ($F_{3,63} = 1.152, p = .335$). At this point, it is unclear whether there is (practically seen) no effect, and designers can use the same thresholds regardless of how people grasp and move the proxy object, or there might be an effect that we could not find due to insensitive data. Therefore, we computed Bayes factors, which for manipulation trajectory 0.918 ($BF_{excl}$) did not favor either hypothesis, and thus, indicates that the data is insensitive [22]. We conclude that more data would be needed to unravel this variable. On the other hand, the $BF_{excl}$ for grasping type is 29.760, suggesting that it is 29.760 times more likely to observe this data under the null hypothesis. Hence, there exists very strong evidence that grasping type did not affect the detection thresholds [42]. For the interaction effect, we found moderate evidence for the null hypothesis ($BF_{excl} = 7.590$).

### 4.1.7 Summary.

The study showed that we can introduce substantial offsets that are undetectable by humans for all grasping types and across both movement trajectories. Our Two-Way Repeated Measures ANOVA could not reveal a main effect in our data, and the Bayes analysis suggests that for movement trajectory, this is due to insensitive data. In contrast, for grasping type, the Bayes analysis provides strong evidence for the null hypothesis, i.e., there exists a high likelihood that grasping type did not affect the detection thresholds.

To this end, we restricted users’ motion to isolate the potential effects of grasping type and movement trajectory on the detection thresholds. However, in our everyday environment, humans regularly perform unrestricted movements with objects, requiring them to lift the objects. Therefore, we conducted a second study to investigate the role of different grasping types and object masses during unrestricted movement.

#### 4.2 Study 2: Effects of Grasping Type and Object Mass

In study 2, we compared four grasping types (lateral, medium wrap, tripod and writing tripod) and two mass conditions: (1) all objects had equal mass and (2) all objects had a range of different masses accordingly to the grasping type. We did not restrict user movements in any way, which introduces some variance. This allows us to isolate the effects that different grasping types and object masses may have on the perception. Through this study, we wanted to understand whether these variables significantly contribute to a semantic violation and therefore require special consideration when designing visuo-haptic illusions.

**4.2.1 Design.** We utilized an adaptive psychophysical 1-up-1-down interleaved staircase procedure with a 4x2 within-subjects design. We had two independent variables: GRASPING TYPE (lateral vs. medium wrap vs. tripod vs. writing tripod) x OBJECT MASS (equal vs. unequal mass). We investigated 8 conditions which were counterbalanced using a Latin square (n=8). We used the same method as in study 1.

**4.2.2 Participants.** We recruited a new set of 24 right-handed participants (nine females, fifteen males), aged 20–37 (mean: 26.70; SD: 5.01) from the general public and the local university. This excludes two participants who were omitted from the analysis due to (1) not reaching convergence in the study and (2) a complete system failure. Participants had a range of different educational and professional backgrounds including computer science, media informatics, electronics, pharmacy, bioinformatics, HCI, mathematics, and psychology. All participants reported normal or corrected-to-normal vision and did not report any known health issues which might impair their perception or proprioception. Two participants had never used VR before, sixteen had used it a few times (one to five times a year), one person used VR often (6–10 times a year), and five others on a regular basis (more than 10 times a year). Nine participants reported that they had not played VR games before, eleven people responded sometimes or infrequently (1–5 times a year), one person used it often (6–10 times a year), and three people on a regular basis (more than 10 times a year). Participants not associated with our institution received €10 as remuneration for taking part in the experiment. The study was approved by the University’s Ethics and Hygiene Board.

**4.2.3 Apparatus.** In study 2, we used the same apparatus as in study 1, but we removed the turntable and the slider. Instead, participants manipulated the proxies directly on the table (see Figure 8). For study 2, we 3D printed eight objects, two of each kind using PLA, and included 3D printed conductive parts to enable touch sensing. The objects were connected via long thin cables to the Arduino Uno, not limiting the interaction space. The design of the lateral object was slightly altered to ensure a more natural manipulation (center of mass). The first set of objects was fabricated to have an equal weight of 40g (± 1g tolerance), whereas the second set was weighted (see Figure 8, right) using lead shot and secured with super glue (± 2g tolerance). The lead shot was equally distributed and superglued inside the objects, providing a realistic center of mass to avoid a
break in immersion [76]. All objects were augmented with three retroreflective markers, allowing us to precisely track them in 3D space using Optitrack. To improve tracking quality and robustness we added paper straws to enable an optimal marker setup. The pivots were carefully calibrated for each object.

4.2.4 Procedure & Data Collection.
We used the same procedure and data collection method as in study 1. Participants were instructed to move the object in the most direct way (linear) to the goal position without dragging it on the table, requiring them to slightly lift the object. This ensured that participants felt the mass of the object.

4.2.5 Results.
We report our estimates for the conservative detection thresholds using different grasping types (lateral, medium wrap, tripod, and writing tripod) with two mass conditions (equal and unequal mass). Then, we analyze the results with respect to our study objectives.

Detection Thresholds for Grasping Types and Object Mass.
Overall, we received 3950 responses in the interleaved-staircase procedure, and it took participants 21.5 (SD: 2.9) trials to reach convergence. As in study 1, each participant contributed eight thresholds, i.e. one per condition, which was determined by averaging the last four reversal points in each sequence. The overall thresholds for all eight study conditions were computed by taking the mean across all 24 individual threshold values. The results can be found in Figure 9. All 192 staircase plots from study 2 are available in the appendix. Similar to study 1, the analyses of the SSQ responses show an increase of the total severity score (mean = 27.43, SD = 22.56).

Analysis.
We further analyzed our data using a Two-Way Repeated Measures ANOVA on the two independent variables grasping type and object mass. There were no extreme outliers in the dataset. A Shapiro–Wilk test indicated a violation of the normality assumption at \( p = .05 \) in the lateral/equal mass condition. Hence, we examined the normal QQ plot (see appendix) and computed skewness \( \gamma_1 \) and kurtosis \( \gamma_2 \) values (\( |\gamma_1| \) and \( |\gamma_2| < 2.3 \)) leading to the conclusion that we can run parametric tests [30, 51]. The dataset met the assumption of sphericity verified through Mauchly’s test at \( p = .05 \). As in study 1, we computed Bayes factors to further analyze our collected data.

Research Objectives R1 & R3. The Two-Way Repeated Measures ANOVA did not reveal a main effect of either variable, object mass (\( F_{1,21} = .371, p = .549 \)) or grasping type (\( F_{3,63} = .430, p = .732 \)). In addition, there was also no interaction effect (\( F_{3,63} = .757, p = .522 \)). Following this analysis, we computed Bayes factors, and for object mass we found moderate evidence (\( BF_{excl} = 5.527 \)) in favor of the absence of an effect on the thresholds, i.e., no effect on the detection thresholds is 5.527 times more likely than that there was an effect. This does not contradict previous findings on the significant impact of force on proprioceptive accuracy. Ansems et al. [3] tested 10%, 25% and 40% of maximum voluntary contraction (MVC) force, which refers to the highest possible load an individual can
move using a muscle (group). In fact, these values are substantially greater than the maximum proxy weight of 500 g in our study. In line with study 1, the Bayes factor for grasping type, 9.623 ($BF_{exc}$), provides moderate evidence for the null hypothesis (R1), and for the interaction effect model, strong evidence ($BF_{exc} = 13.112$) for the null hypothesis given the data.

### 4.2.6 Summary

In this section we reported our estimates for the conservative detection thresholds. The Two-Way Repeated Measures ANOVA did not show a main effect. However, the Bayes analysis provides evidence for accepting the null hypothesis on the variables grasping type and object mass – there exists a high likelihood that neither variable, grasping type or object masses (≤ 500 g), affected the amount of disparity which can be introduced. Next, we analyze both studies with respect to our study objectives R4 and R5.

### 4.3 Restricted vs. Unrestricted Linear Movement & Individual Differences

In this section we analyze both studies, in total 48 participants contributing 384 thresholds, to investigate the differences between linear restricted vs. unrestricted movement type (R4). We observed a high threshold variance across participants which led to the question whether there are consistent differences in humans’ proprioceptive acuity (R5). Finally, we analyze participants’ backgrounds with respect to the determined thresholds to better understand where such differences may come from.

#### 4.3.1 Restricted vs. Unrestricted Linear Movement

Here, we make the assumption that grasping type did not have an effect on the detection thresholds following the evidence obtained through our Bayes analysis in both studies. We analyzed the between-subjects factor linear movement type (restricted vs. unrestricted movement). Since there are two levels in the unrestricted movement condition (equal weight and weighted), we ran two independent samples Welch’s t-tests, because the dataset did not meet the homogeneity of variance assumption verified through Levene’s test. A Shapiro–Wilk test indicated a violation of the normality assumption at $p = .05$. Given our sample size, we examined the normal QQ plot (see appendix) and computed skewness $\hat{\gamma}_1$ and kurtosis $\hat{\gamma}_2$ values (for all conditions $|\hat{\gamma}_1| \text{ and } |\hat{\gamma}_2| < 3$) leading to the conclusion that there is no severe violation of normality [30, 51]. Further, we applied Bonferroni corrections to account for Type I errors. Additionally, we performed two Bayesian independent samples t-tests using default effect size priors. Results are reported as two-tailed Bayes factors $BF_{10}$ and effect size estimates as median posterior Cohen’s $\delta$ with a 95% credibility interval (95%CI) [46].

#### Research Objective R4

Our analysis provides strong evidence for an increase in detection thresholds in the linear restricted movement condition (Mdn = 1.45), when comparing to both unrestricted conditions: (1) linear weighted (Mdn = 1.33) ($t_{(180)} = 4.13, p < .001, d = 0.596, BF_{10} = 350.080$, with median posterior $\delta = [0.282, 0.855]$), and (2) linear equal weight (Mdn = 1.34) ($t_{(180)} = -4.36, p < .001, d = -0.629, BF_{10} = 826.248$, with median posterior $\delta = -0.600, 95$%CI = [-0.890, -0.313]). These results suggest that limiting a user’s DoF reduces proprioceptive accuracy and thus allows for greater discrepancy (see Figure 10). Based on our observations, we believe that this is caused by the ‘somewhat’ artificial movement and the momentum that is generated when smoothly manipulating the object along a fixed trajectory (slider/turntable). In contrast, unrestricted linear manipulations resemble a frequently occurring, highly trained and memorized interaction which may lead to higher accuracy.

#### 4.3.2 Proprioceptive Differences

As illustrated in Figures 8, 9 and 10, participants’ thresholds are widely spread across the entire testing spectrum, which leads to the question: *Are humans equally sensitive to visuo-haptic illusions?* We ran a descriptive analysis on our study 1 and study 2 data. Then, we computed an overall threshold for each participant by averaging their 8 detection thresholds.

#### Research Objective R5

The datasets follow a normal distribution with a consistent SD, indicating that it is representative for the general population. Since there are no additional density humps (which would suggest multiple performance groups) we conclude that all participants belong to the same population (see appendix). However, we were still surprised by how large the threshold spectrum is, reaching from 5% to almost 67% possible C/D gains. Therefore, we analyzed whether these differences are connected to participants’ backgrounds reported in the demographic questionnaires.
**Analysis.** We conducted multiple Spearman’s ρ rank correlations across all 48 participants, evaluating if there is a relationship between the detection thresholds and the ratings on the following questionnaire items: participating in physical sports activities, prior VR experience, prior experience with 3D interactions, experience with VR gaming, gender, and age.

**Results.** There was a positive correlation between the two variables threshold and prior VR experience (r(46) = .15, p = .003), indicating that with more VR experience, thresholds become smaller. There was no correlation between participants’ thresholds and physical activities (r(46) = .03, p = .567), experience with 3D interactions (r(46) = .03, p = .502), VR gaming (r(46) = -.04, p = .385), gender (r(46) = -.03, p = .592) or age (r(46) = -.02, p = .682).

4.3.3 Summary.
We found that restricting a user’s movement results in significantly higher thresholds. Additionally, we investigated where differences in proprioceptive accuracy may be linked to. It appears that one of the important factors is previous experience in VR, impacting how much discrepancy can be introduced. In the next section we outline how our results support designers when incorporating visuo-haptic illusions in their workflow.

5 DESIGNING VISUO-HAPTIC ILLUSIONS
Here, we demonstrate the impact of our key findings on the design process of visuo-haptic illusions in VR. For illustration purposes, we use the cooking class scenario in which our VR designer Lisa wants to include visuo-haptic illusions to enhance the experience (see Figure 11). Lisa is thinking about the possible interactions with objects and the environment as well as the physical proxies that are available, and encounters the challenges below:

**Grasping Types (A).** Lisa is designing the kitchen space consisting of a virtual stove and a single proxy which acts as a stand-in for all the available virtual pots (and maybe pans and bowls) in the kitchen. By using visuo-haptic illusions she can simulate different sized pots on the stove. Consider that there are several ways that kitchen utensils such as a whisk could be grasped. Here, she needs to be aware whether the differences in how the tool is handled affect the extent to which an illusion can be used. Following our results, she would not have to restrict the interaction type in any way. Hence, a user can seamlessly transition between different grasping types, resulting in a natural and realistic experience.

**Proxy Mass (B).** Depending on the desired dish, the students need different kitchen utensils (e.g., a whisk, a hand mixer, or spoons). Clearly, each tool is suited for a specific use case; thus, they differ in their properties and dimensions. Given these uncertainties, Lisa needs to understand which attributes might limit or expand the illusion space. Our results suggest that for handheld sized objects (≤ 500 g), the amount of movement discrepancy which can be introduced is not noticeably affected by the object’s mass.

**Restricted vs. Unrestricted Movement (C).** Finally, we demonstrate how increased detection thresholds may help to enhance proxy-based interactions. For instance, a student is holding a virtual baking sheet embodied by a physical (document holder) proxy. The physical proxy does not perfectly match the depth of the virtual oven. By using a visuo-haptic illusion, Lisa can create a “matching” depth sensation. She is aware that the oven rails restrict a user’s movement; this allows her to include a larger offset, which would otherwise not have been possible.

6 DISCUSSION & FUTURE WORK
Our work aims to untangle the contributing variables leading to a semantic violation when manipulating a virtual object embodied by a physical proxy. We believe that our results can be seen as a first step towards a generalizable approach for visuo-haptic illusion design. Here, we discuss our findings in a broader context, outline current limitations, and give recommendations for future work.

6.1 Role of Movement Speed
Aligned with other researchers’ observations [26], in our pilot studies we observed that movement speed seems to be a critical variable. To the best of our ability, we tried to control for it by (1) instructing participants to move the object with a consistent “normal” speed, (2) giving them a warm-up round to establish a comfortable pacing and (3) monitoring their speed through our study program. As soon as participants moved the object faster or slower than the previously determined threshold boundaries, they were instructed (by audio) to adjust their speed. Future work should aim to investigate the role of movement speed in visuo-proprioceptive conflicts.

6.2 Restricted vs. Unrestricted Movement
The potential differences between restricting and not restricting the DoF of a user’s motion has powerful implications for the design process of visuo-haptic illusions. In fact, many real-world objects,
UI elements and mechanics limit or guide users’ movements, such as steering wheels, levers, switches, shifters, door handles, keyholes, sliders, knobs and many others. Capitalizing on this potential when designing VR illusions can help to create more engaging, realistic, and powerful applications. Moreover, there are several haptic devices which steer or guide a user’s movement, such as ElastiLinks [72] and Haptic Links [69], which could greatly benefit from illusions. To this end, our results build on the assumption that different grasping types do not have an effect on the thresholds. Therefore, future work should focus on direct comparison between restricted vs. unrestricted movements and additionally, incorporate other trajectories which we did not compare.

6.3 Generalizability

It is important to note that we only investigated gain factors $\geq 1.0$, scaling up a user’s real-world movement which is in practical terms more prevalent. Nevertheless, at this point it is unclear whether scaling down (gain factors $< 1.0$) may reveal other results. Further, we chose our set of diverse grasping types and object masses based on the selection criteria described above, covering a wide spectrum to detect potential differences; however, there is still the possibility that other grasping types/object masses might influence the amount of discrepancy which can be introduced. In addition, the role of movement trajectory needs further investigation, because our study could not unpack it. Hence, future studies are needed to gather more evidence leading to a better understanding of the relevant variables in visuo-proprioceptive conflicts.

Finally, an interesting question for future work is how the visualization (i.e., only showing a user’s hand and not the entire arm chain, thus offering very limited visual cues to detect a manipulation) generally affects the detection thresholds.

6.4 Personalized VR Experiences

To our surprise, we found that individual thresholds differ quite drastically. Our analysis suggests that this is linked to participants’ previous experience in VR. Likewise, there are many more variables contributing to an individual’s proprioceptive acuity which we could not assess in our questionnaire [58]. Since the perceptual differences appear to be widely spread, we propose to investigate whether we can establish a method for proprioceptive calibration. In fact, this could also be expanded to other illusion techniques such as redirected walking [66], haptic retargeting [5] or redirected touch [48]. This approach could first calibrate a conservative base threshold, and depending on other parameters such as complexity of the experience (distracting factors) [16, 23, 77] or time spent in the application [58], we could constantly adjust the illusion, delivering a VR experience tailored to the individual.

7 CONCLUSION

In this work we unravel the extent to which three variables (grasping type, movement trajectory and object mass) impact the amount of discrepancy which can be introduced while remaining unnoticed by a user. Our frequentist analysis did not reveal a significant effect of the study variables grasping types, object masses ($\leq 500$ g) and the different movement conditions (linear and circular) on the amount of discrepancy which can be introduced. The computed Bayes factors suggest that for movement trajectory, this was due to insensitive data. For grasping type and object mass, there was evidence for the absence of an effect on the detection thresholds. However, we found a significant effect between linear restricted and unrestricted movement. Restricted movement led to smaller detection thresholds, indicating that proxy manipulations which limit a user’s motion along a fixed path may allow for greater discrepancy. We identified a wide range of thresholds linked to participants’ prior experience in VR, suggesting that we need some sort of proprioceptive calibration process — pushing towards personalized VR experiences. Finally, we outlined a design scenario demonstrating the relevance of our results in a practical context.

ACKNOWLEDGMENTS

We would like to thank Denise Türk for helping with visuals and Anna Milena Rothermel, Nina Knieriem and Rudolf Siegel for their advice on the analysis. This research was funded in part by the German Federal Ministry of Education and Research (BMBF) under grant number 01IS17043 (project VRGIO) and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project numbers 425868555 and 450247716.
Designing Visuo-Haptic Illusions with Proxies in Virtual Reality


